PRECISION MEASUREMENT OF THE HIGGS BOSON MASS AND SEARCH FOR DILEPTON MASS RESONANCES IN H $\rightarrow 4\ell$ DECAYS USING THE CMS DETECTOR AT THE LHC

By

JAKE ROSENZWEIG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2022

I dedicate this to Jacob Myhre.

ACKNOWLEDGEMENTS

Without so many two- and four-legged blessings along the way, I could never have made it to this point in my academic career. Thus, I give my infinite thanks to the following folks.

To my high energy physics mentors, Professors Andrey Korytov and Guenkah Mitselmakher, for granting me this one-of-a-kind opportunity to do real *science* at CERN.

To my wife, Suzanne Rosenzweig, for showing me that dreams do come true. To my mother and father, Vicki and John, who always reassured me that I could achieve anything I put my mind to. Sleep peacefully, Mom. To my siblings, Alex, Ryan, Devin, Jace, and Claudia who frequently and gently reminded me that there was life outside of grad school.

Aunt Rach, Uncle Yuri,

To my mentor, Sheldon Friedman, and his wife, Rita Friedman (Rosenzweig), who chose to invest in my success at a young age. I have only made it this far thanks to your undying encouragement, love, and optimism. To Sheldon's best friend, Dr. Bernard Khoury, whose reputation has helped pave my road.

To Dr. Filippo Errico for his Dr. Lucien Lo To Dr. Noah Steinberg To Darin Acosta, for spending hours of discussion on every t To the gentle gents who introduced me to the world of CMS, Brendan Regnery and Bhargav Joshi.

To my comrades for showing me what it takes to survive the core courses, Dr. Atool Divakarla, Dr. Brien O'Brendan, Dr. Donyell Guerrero, and Dr. Vladinar Martinez.

To the many students who tagged along in our "CMS Office Hours": Sean Kent, Jeremiah Anglin, Cris Caballeros, Ari Gonzalez, Evan Koenig, Nik Menendez, Neha Rawal, John Rötter. And to the many students who let us practice our spiels:

To my mentee, Matthew Dittrich, for

To my Polish roommates in Saint-Genis-Pouilly for showing me what home away from home feels like.

To the many moms who generously gave unconditional support during the darkest times and unconditional love during the brightest times: Silet Wiley, Margaret Sherrill, Dawn Hood, Cyndi Reilly-Rogers.

To my childhood best friends: Jish, Willis, Shane, Zac, Duck, and Marcus for their constant

clever competition which has shaped me into the determined man I am today.

Big Tree:

And finally to Existence itself for this unpredictable, unbelievable blip of an experience called life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figures page

CHAPTER 1
HIGGS BOSON MASS MEASUREMENT IN THE $H \rightarrow ZZ^{*} \rightarrow 4\ell$ CHANNEL

1.1 Motivation

The Higgs boson was discovered in 2012 by the CMS and ATLAS collaborations. This was a momentous achievement in particle physics because the existence of the Higgs boson was required to complete the SM. In fact, it is sometimes referred to as the “missing puzzle piece” of the SM. The Higgs boson is one of a kind: it is the only fundamental scalar particle ever discovered so far. The unique boson could be a portal to new physics (*beyond Standard Model physics*, BSM), e.g., by decaying into BSM low-mass dilepton mass resonances (Chapter **??**). In order to be certain that the recently discovered Higgs boson is truly the same as the one predicted by the SM, it is necessary to compare its measured properties to the predicted ones.

Some properties of the Higgs boson can be predicted by the SM, like - There are many results on Higgs properties: spin, charge, decay processes, lifetime, mass. - The last of these is the focus of this dissertation and is of particular importance to the Universe: depending on mH and mtop, the stability of the Universe.

ALL previous mass measurements: - Run 1: - H->2gamma VALUE - H->ZZ->4L VALUE - Run 2: - H->2gamma VALUE - (2016) H->ZZ->4L VALUE - H->bb - H->mumu - H->WW

- Why this thesis is important: - This thesis describes the methodology and results of the best precision measurement of mH to date by using the hZZ4l decay and Full Run 2 data set from CMS. - Run 2 provides more data -> more precision on measurements of Higgs properties. - In addition to more HZZ4l events, this analysis provides new techniques, specifically the VX constraint. - Predict mH for Run 3, will start soon summer 2022 and provide an approximate 300? /fb of L int. - In 2026(?), HLLHC provides even more data. ref snowmass paper.

This chapter is structured as follows:

- Section 1.2: General overview of the analysis of the Higgs boson mass measurement.
- Section 1.3: Data sets, triggers, and simulation.
- Section 1.4: Event reconstruction and selection.
- Section **??**: Background estimation.
- Section 1.5: Signal modeling and improvements, including kinematic discriminant, per-event mass uncertainties, VXBS constraint, reference to ad hoc studies in appendix.

- Section 1.7: Systematic uncertainties.
- Section 1.8: Results.
- Section 1.9: Summary.

1.2 Analysis Overview

The first step to performing a precision measurement of the Higgs boson mass (m_{H}) is to "observe" many Higgs bosons. However, production of a Higgs boson is essentially nonexistent in everyday conditions and is still extremely rare even in the high-energy pp collisions of the LHC. At a center-of-mass energy of 13 TeV, the total inclusive inelastic cross section of two protons colliding is 70 mb TODO: CITE. Comparing this to the production cross section of a Higgs boson (TODO sigma(pptoH) = 59 pb) shows that a Higgs boson is produced in approximately one out of every billion pp collisions. TODO CITE

To complicate matters further, the Higgs boson has a *very* short mean lifetime of only 1.6×10^{-22} s [1]. Thus, the Higgs boson is not directly detected by CMS but is instead *inferred* from its stable decay products that enter the various subdetectors. Among all the fundamental particles so far discovered, the Higgs boson bears the second heaviest mass (approximately 125 GeV), the first belonging to the top quark (Section **??**). This gives the scalar boson sufficient energy to decay into at least 9 different final states. MENTION THAT NOT ALL DECAYS MAKE ON-SHELL PARTICLES? Each decay occurs with a different probability—the *branching fraction* or *branching ratio* ($\mathcal{B}$)—whose value depends on m_{H} as shown in Figure 1-1. The question then becomes, *"Which decay mode of the Higgs boson is most useful for the measurement of m_{H}?"*.

Owing to its large signal-to-background ratio of approximately 2 and its relatively rare four-lepton final state, the $\mathrm{H} \to \mathrm{ZZ}^{*} \to 4\ell$ decay channel is selected and is called the *signal* process. Thus, a single Higgs boson will decay via the signal process into two Z bosons (one on-shell and one off-shell) on average only 2.6% of the time. In turn, each Z boson decays into two opposite-sign, same flavor (OSSF) leptons ($\mathrm{Z} \to \ell^{+}\ell^{-}$, where $\ell = \mathrm{e}, \mu$) on average approximately 6.7% of the time, giving rise to four distinct final states: 4e, 4μ, $2\mathrm{e}2\mu$, $2\mu 2\mathrm{e}$. The

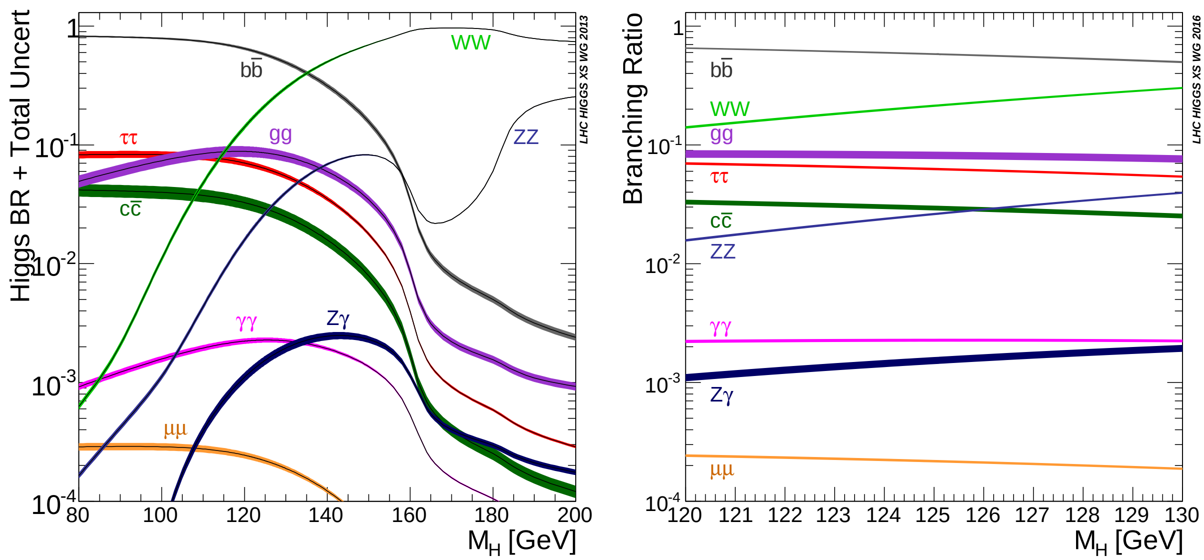

Figure 1-1. The branching ratios of various Higgs boson decays as a function of the Higgs boson mass over a wide range (Left) and a narrow range (Right) of values.

branching ratio for the overall signal process is then calculated as:

$$\mathcal{B}\left(\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell\right) = \mathcal{B}\left(\mathrm{H} \to \mathrm{ZZ}^*\right)\left[\mathcal{B}\left(\mathrm{Z} \to \ell^+\ell^-\right)\right]^2 = 1.8 \times 10^{-3}.$$

Thus, a signal event is expected to be produced only once in about every *trillion* pp collisions.

The strategy is then to search the pp collision data collected and analyzed by the CMS detector (Chapter **??**) for all the detected $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ events. The task is not so straightforward; events in the data are categorized—not by the entire decay process—but by their final state, based on which triggers fired to collect which events. Section 1.3 describes the triggers used for this analysis to select events with the 4ℓ final state found in the corresponding data sets. For each chosen event, the subdetectors of CMS (Chapter **??**) provide a plethora of track and energy-detection information to reconstruct *objects*—representations of the underlying particles within the event. The reconstructed objects are then assembled in a fashion that checks if the logic coincides with the process of interest: $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$. For example, a pair of OSSF lepton-like objects should appear to come from a Z-like object—i.e. having a nominal mass of approximately 91 GeV and zero net electric charge—instead of, say, appearing to come from a H-like object. Two such Z-like objects must be formed and should appear to come from a H-like object. All throughout, the reconstructed event must obey physics conservation laws (energy, momentum, charge, etc.) and the associated objects may even be required to pass certain detector selection criteria (e.g. $p_{\mathrm{T}}^{\mu} > 5\,\mathrm{GeV}$). These criteria are analysis-specific and are collectively called the *event selection* of the analysis. The event selection for this analysis is described in Section 1.4.

Even though the event selection is constructed to select signal events, it is not guaranteed; there are certain physics process that have exactly the same initial and final states as the signal process. Such processes "contaminate" the collected signal events and are called *background processes*. Figure **??** shows how identical initial state gluons can react to produce exactly the same final state particles, while producing different intermediate particles: the signal process (Left), initiated by gluon-gluon fusion vs. a background process (Right) which skips the intermediate Higgs boson. It is imperative for all physics analyses to maximize the number of collected signal events while minimizing the number of collected background events. Section **??** discusses the associated background processes and how to estimate the number of events these contribute to the signal region.

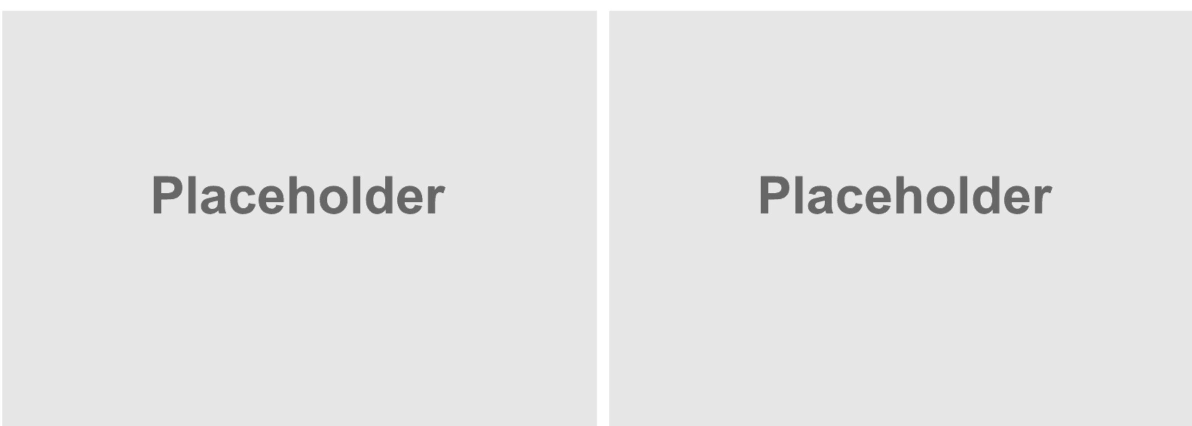

Figure 1-2. Feynman diagrams showing how the initial and final states are the same for the signal process ($gg \to H \to ZZ^* \to 4\ell$, Left) and one possible background process ($gg \to ZZ \to 4\ell$, Right).

Before drawing conclusions from the data themselves, it is necessary for particle physicists to make predictions about their analysis using simulated samples or *simulation*. These samples contain simulated events usually of a specific process (e.g. $\mathrm{pp} \rightarrow \mathrm{H} \rightarrow \mathrm{ZZ}^* \rightarrow 4\ell$), governed by some theoretical framework that is programmed mathematically into the software package. Programs like MadGraph5_amc@nlo and powheg can simulate millions of rare (or even *fictitious*) events, which might otherwise take many years to observe in data. Furthermore, software can even simulate the particles as they travel through the simulated detectors. Programs like Geant4 can show analysts what to expect as the particles interact with a virtual version of the CMS detector. Predictions from simulation can then be compared to the truth—the data—as a way to check the accuracy of the analysis. For example, a surplus of events in data where none was expected may lead to the discovery of new particles, as was the case in the discovery of the Higgs boson. The simulated samples for this analysis are described in Section 1.3.

So how is the measurement of m_{H} achieved? Since the signal process is $\mathrm{H} \rightarrow \mathrm{ZZ}^* \rightarrow 4\ell$, conservation of energy leads one to expect that $m_{4\ell} \approx m_{\mathrm{H}}$. Although this is not how the final measurement is obtained, it is a logical starting point. The distribution of $m_{4\ell}$ values reveals that the Higgs boson mass resonance stands well above the distribution of expected background events (Figure 1-3). Simulation is then used to predict the *line shape* of this signal peak (Section 1.5). This signal modeling is performed using a double-sided Crystal Ball function to fit the line shape, for various mass points of m_{H}, in each of the four final states.

The precision of the measurement of m_{H} is improved by implementing several techniques: calculating a per-event matrix element kinematic discriminant, deriving correction factors for $m_{4\ell}$ uncertainty for various regions of phase space, reevaluating the lepton p_{T} values using a Z_1 mass constraint, and constraining the four muon tracks to the selected vertex (*vertex constraint*). The implementation of these techniques is discussed in Section 1.6.

Important in any scientific analysis is the careful study of all the associated uncertainties that inherently come with making *any* measurement. The two main kinds of uncertainties are either *statistical*, which depend on the number of data points used to the make the measurement, or *systematic*, which depend on the instrumentation precision used to make the measurement. The systematic uncertainties for this analysis include:

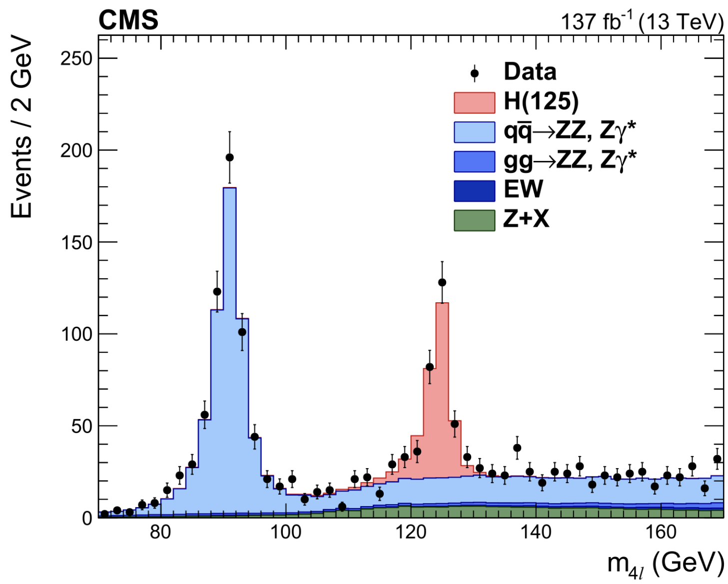

Figure 1-3. Distribution of $m_{4\ell}$ from $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ events using Full Run 2 data.

- Lint (TODO 2.5%)
- Lepton identification and reconstruction efficiencies (TODO 2.5–9%): lepton energy scale
- BOTH OF THE ABOVE AFFECT SIGNAL AND BKG
- Estimation of the reducible background.

The systematic theoretical uncertainties include:

- renormalization and factorization scale uncertainties
- choice of the set of parton distribution functions
- BOTH OF THE ABOVE AFFECT SIGNAL AND BKG
- branching fraction uncertainties for signal and background processes.

These uncertainties are discussed in Section 1.7.

The resolution of the peak Ways to improve the Do likelihood fit. $\mathcal{L}$

1.3 Data Sets, Simulated Samples, and Triggers

1.3.1 Triggers

1.3.2 Data Sets

1.3.3 Simulated Samples

1.4 Event Reconstruction and Selection

1.4.1 Event Reconstruction

1.4.2 Event Selection

1.4.2.1 ZZ Candidate Selection

1.5 Signal Modeling

1.5.1 Z_1 Mass Constraint

1.5.2 Per-Event Relative Mass Error Categorization

1.5.3 VXBS Constraint

1.5.4 Matrix Element-Based Kinematic Discriminant

Expected uncertainty	4μ	4e	$2e2\mu$	$2\mu2e$	Inclusive
1D (No bkg)	153	466	315	300	121

Table 1-1

1.6 Techniques to Improve Mass Resolution

Words!

1.7 Systematic Uncertainties

Words.

1.8 Results

Words.

1.9 Summary

Words.

REFERENCES

[1] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020. doi: 10.1093/ptep/ptaa104. and 2021 update.